Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON QUADRATIC FORMS

BY PAUL SAUREL

IN a recent note* we have reproduced the very simple demonstration by means of which Gibbs establishes the conditions for a positive quadratic form. It may be of interest to show that the more general question of determining the number of positive and of negative terms in the expression of a quadratic function as a sum of squares can be treated in a similar manner.

Let us consider the quadratic function φ defined by the equation

$$\varphi = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} x_i x_k , \tag{1}$$

in which

$$a_{ik} = a_{ki}, \tag{2}$$

and let us write

$$f_i = \sum_{k=1}^{n} a_{ik} x_k . \tag{3}$$

From equations 1 and 3 it follows that

$$\varphi = \sum_{i=1}^{n} f_i x_i . \tag{4}$$

If $a_{11} \neq 0$ we can throw this equation into the form

$$\varphi = \frac{f_1^2}{a_{11}} + \sum_{i=2}^{n} f_i^{(1)} x_i , \tag{5}$$

where

$$f_i^{(1)} = \frac{a_{11} f_i - a_{1i} f_1}{a_{11}} . \tag{6}$$

Since $f_i^{(1)}$ is independent of x_1, it follows that the second of the two terms on the right-hand side of equation 5 is independent of x_1. Moreover, it should be noticed that the coefficient of f_1^2 is the reciprocal of

$$\left(\frac{\partial f_1}{\partial x_1}\right)_{x_1,\, x_2,\, x_3,\, \ldots\, ,\, x_n} , \tag{7}$$

---

* ANNALS OF MATHEMATICS, vol. 4 (1902-03), p. 62.

in which the subscripts indicate the quantities which are regarded as independent variables.

If a_{11} be equal to zero, equations 5 and 6 are meaningless. We can, however, obtain a pair of analogous equations by employing any one of the functions f_i for which the corresponding coefficient a_{ii} is not equal to zero. Then by changing the numbering of the variables we can again obtain equations 5 and 6.

We have accordingly the following theorem:

Theorem I. *If there be one of the quantities*

$$\left(\frac{\partial f_i}{\partial x_i}\right)_{x_1, x_2, x_3, \ldots, x_n}$$

which is not equal to zero, the given quadratic function of n variables can be expressed as the sum of a quadratic function of not more than n—1 variables and the square of a linear function. The coefficient of this last term has the sign of the non-vanishing derivative.

If

$$a_{11} = a_{22} = \cdots = a_{nn} = 0, \tag{8}$$

the transformation given above cannot be performed. In this case we shall write

$$\varphi = \frac{2f_1 f_2}{a_{12}} + \sum_{i=3}^{n} f_i^{(2)} x_i = \frac{(f_1 + f_2)^2 - (f_1 - f_2)^2}{2a_{12}} + \sum_{i=3}^{n} f_i^{(2)} x_i, \tag{9}$$

where

$$f_i^{(2)} = \frac{a_{12} f_i - a_{1i} f_2 - a_{2i} f_1}{a_{12}}. \tag{10}$$

Since $f_i^{(2)}$ is independent of x_1 and x_2, it follows that the last term on the right hand side of equation 9 is also independent of x_1 and x_2. If a_{12} be equal to zero the above transformation becomes impossible. In this case, however, an analogous transformation can be performed with a pair of functions f_1, f_i for which the corresponding coefficient a_{1i} is not equal to zero. Such a coefficient exists; otherwise, the given function of n variables would be independent of x_1. The variables can now be renumbered so that equations 9 and 10 are again obtained.

We have accordingly the following theorem:

THEOREM II. *If each of the quantities*

$$\left(\frac{\partial f_i}{\partial x_i}\right)_{x_1,\, x_2,\, x_3,\, \ldots,\, x_n}$$

be equal to zero, the given quadratic function of n variables can be expressed as the sum of a quadratic function of not more than n—2 variables and the squares of two linear functions. The coefficients of these last two terms have opposite signs.

It is evident that repeated applications of one or both of the transformations just described will yield the following theorem:

THEOREM III. *A quadratic function of n variables can be expressed as the sum of the squares of not more than n linear functions.*

Let us suppose that by repeated applications of Theorems I and II the given quadratic function has been expressed as the sum of r squares and a quadratic function which is independent of $x_1, x_2, \cdots, x_r$. We shall suppose that this quadratic function has been written in each of the forms:

$$\sum_{i=r+1}^{n} f_i^{(r)} x_i, \tag{11}$$

$$\sum_{i=r+1}^{n} \sum_{k=r+1}^{n} a_{ik}^{(r)} x_i x_k \,. \tag{12}$$

If the last transformation employed be one of the type 5, $f_i^{(r)}$ is defined by the following equation, analogous to equation 6:

$$f_i^{(r)} = \frac{a_{rr}^{(r-1)} f^{(r-1)} - a_{ri}^{(r-1)} f_r^{(r-1)}}{a_{rr}^{(r-1)}} \,. \tag{13}$$

If, on the other hand, the last transformation employed be one of the type 9, $f_i^{(r)}$ is defined by the following equation, analogous to equation 10:

$$f_i^{(r)} = \frac{a_{r-1,r}^{(r-2)} f_i^{(r-2)} - a_{r-1,i}^{(r-2)} f_r^{(r-2)} - a_{r,i}^{(r-2)} f_{r-1}^{(r-2)}}{a_{r-1,r}^{(r-2)}} \,. \tag{14}$$

In either case

$$a_{ik}^{(r)} = \left(\frac{\partial f_i^{(r)}}{\partial x_k}\right)_{x_{r+1},\, x_{r+2},\, \ldots,\, x_n} \,. \tag{15}$$

We shall now show that

$$a_{ik}^{(r)} = \left(\frac{\partial f_i}{\partial x_k}\right)_{f_1,\, f_2,\, \ldots,\, f_r,\, x_{r+1},\, \ldots,\, x_n} \,. \tag{16}$$

Consider the expression

$$\left(\frac{\partial f_i^{(1)}}{\partial x_k}\right)_{x_2,\, x_3,\, \ldots\, ,\, x_n} \tag{17}$$

Since $f_i^{(1)}$ is independent of x_1 we may give to dx_1 any convenient value. It will therefore be allowable to suppose that dx_1 has been so taken that

$$df_1 = 0. \tag{18}$$

But, in that case, equation 6 shows that

$$df_i^{(1)} = df_i. \tag{19}$$

The expression 17 is therefore equal to

$$\left(\frac{\partial f_i}{\partial x_k}\right)_{f_1,\, x_2,\, x_3,\, \ldots\, ,\, x_n} . \tag{20}$$

Consider next the expression

$$\left(\frac{\partial f_i^{(2)}}{\partial x_k}\right)_{x_3,\, x_4,\, \ldots\, ,\, x_n} , \tag{21}$$

in which $f_i^{(2)}$ is defined by equation 10. Since $f_i^{(2)}$ is independent of x_1 and x_2 we may give to dx_1 and dx_2 any convenient values. It will therefore be allowable to suppose that dx_2 has been so taken that

$$df_1 = 0, \tag{22}$$

and that dx_1 has been so taken that

$$df_2 = 0. \tag{23}$$

But, in that case, equation 10 shows that

$$df_i^{(2)} = df_i. \tag{24}$$

The expression 21 is therefore equal to

$$\left(\frac{\partial f_i}{\partial x_k}\right)_{f_1,\, f_2,\, x_3,\, \ldots\, ,\, x_n} \tag{25}$$

In a similar manner we can show that

$$\left(\frac{\partial f_i^{(p)}}{\partial x_k}\right)_{f_{p+1}^{(p)},\, f_{p+2}^{(p)},\, \ldots\, ,\, f_r^{(p)},\, x_{r+1},\, \ldots\, ,\, x_n} \qquad i,\, k > r, \tag{26}$$

is equal to

$$\left(\frac{\partial f_i^{(p-1)}}{\partial x_k}\right)_{f_p^{(p-1)}, f_{p+1}^{(p-1)}, f_{p+2}^{(p-1)}, \ldots, f_r^{(p-1)}, x_{r+1}, \ldots, x_n} \qquad i, k > r, \qquad (27)$$

or to

$$\left(\frac{\partial f_i^{(p-2)}}{\partial x_k}\right)_{f_{p-1}^{(p-2)}, f_p^{(p-2)}, f_{p+1}^{(p-2)}, f_{p+2}^{(p-2)}, \ldots, f_r^{(p-2)}, x_{r+1}, \ldots, x_n} \qquad i, k > r, \qquad (28)$$

according as $f_i^{(p)}$ is defined by an equation of the form 13 or by one of the form 14.

In the first case, since $f_i^{(p)}$ is independent of x_p, we may give to dx_p any convenient value. It will therefore be allowable to suppose that dx_p has been so taken that

$$df_p^{(p-1)} = 0. \qquad (29)$$

But then equation 13 shows that

$$df_i^{(p)} = df_i^{(p-1)}, \qquad i > p, \qquad (30)$$

and expression 26 transforms at once into expression 27.

In the second case, since $f_i^{(p)}$ is independent of x_{p-1} and x_p, we may give to dx_{p-1} and dx_p any convenient values. It will therefore be allowable to suppose that dx_p has been so taken that

$$df_{p-1}^{(p-2)} = 0, \qquad (31)$$

and that dx_{p-1} has been so taken that

$$df_p^{(p-2)} = 0. \qquad (32)$$

But then equation 14 shows that

$$df_i^{(p)} = df_i^{(p-2)}, \qquad i > p, \qquad (33)$$

and expression 26 transforms at once into expression 28.

It now follows immediately that, by repeated applications of one or both of the equalities just established, equation 15 can be transformed into equation 16.

Equation 16 taken in connection with Theorems I and II enables us to state the following fundamental theorem:

THEOREM IV. *Let us suppose the variables to be so numbered that, by means of the two transformations described above,* $x_1, x_2, \ldots, x_r$ *can be successively eliminated from the given quadratic function, one or two at a time, and let us further suppose that every one of the derivatives*

$$\left(\frac{\partial f_i}{\partial x_k}\right)_{f_1, f_2, \ldots, f_r, x_{r+1}, \ldots, x_n} \qquad i, k > r \qquad (34)$$

is equal to zero. Then the given quadratic function can be expressed as the sum of r squares. The signs of these r squares are the same as the signs of the following r derivatives:

$$\begin{aligned}
&\left(\frac{\partial f_1}{\partial x_1}\right)_{x_1, x_2, \ldots, x_n}, \\
&\left(\frac{\partial f_2}{\partial x_2}\right)_{f_1, x_2, \ldots, x_n}, \\
&\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&\left(\frac{\partial f_r}{\partial x_r}\right)_{f_1, f_2, \ldots, f_{r-1}, x_r, \ldots, x_n}.
\end{aligned} \tag{35}$$

provided none of these derivatives be equal to zero. If, however, one of these derivatives be equal to zero, we shall strike it and the following derivative from the list; for every pair of derivatives thus omitted we shall have one positive and one negative square.

The derivatives which appear in the statement of this theorem can be put into another form. Denote by Δ_r the determinant

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix}.$$

If we make use of equations 3 we obtain without difficulty the following equations:

$$\begin{aligned}
&\left(\frac{\partial f_1}{\partial x_1}\right)_{x_1, x_2, \ldots, x_n} = \Delta_1, \\
&\left(\frac{\partial f_2}{\partial x_2}\right)_{f_1, x_2, \ldots, x_n} = \frac{\Delta_2}{\Delta_1}, \\
&\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&\left(\frac{\partial f_r}{\partial x_r}\right)_{f_1, f_2, \ldots, f_{r-1}, x_r, \ldots, x_n} = \frac{\Delta_r}{\Delta_{r-1}}.
\end{aligned} \tag{36}$$

We obtain also the equation

$$\left(\frac{\partial f_i}{\partial x_k}\right)_{f_1, f_2, \ldots, f_r, x_{r+1}, \ldots, x_n} = \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2r} & a_{2k} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ a_{r1} & a_{r2} & \cdots & a_{rr} & a_{rk} \\ a_{i1} & a_{i2} & \cdots & a_{ir} & a_{ik} \end{vmatrix}}{\Delta_r}, \quad i, k > r. \tag{37}$$

If one of the derivatives in 35, the pth for example, be equal to zero, we shall have from equations 36

$$\Delta_p = 0.$$

Moreover, since it was possible to choose as independent variables $f_1, f_2, \ldots, f_{p-1}, x_p, \ldots, x_n$, it follows from equations 3 that

$$\Delta_{p-1} \neq 0.$$

If the derivative on the left side of equation 37 be equal to zero, it follows that the numerator of the right side must be equal to zero. Moreover, since it was possible to choose $f_1, f_2, \cdots, f_r, x_{r+1}, \cdots, x_n$ as independent variables, it follows from equations 3 that

$$\Delta_r \neq 0.$$

Finally, it is well known from the theory of linear equations* that, if the determinant in the numerator of the right side of equation 37 be equal to zero for all values of i and k, then every minor of order $r+1$, and consequently every minor of higher order, that can be formed from the determinant Δ_n will be equal to zero.

These results enable us to restate Theorem IV in the following well-known form:†

THEOREM V. *Consider the series of terms*

$$1, \Delta_1, \Delta_2, \cdots, \Delta_r, \cdots, \Delta_n.$$

It is possible to number the variables so that

$$\Delta_{r+1} = \Delta_{r+2} = \cdots = \Delta_n = 0, \tag{38}$$

---

* *Cf.* Weber, *Lehrbuch der Algebra*, vol. 1, p. 103.

† *Cf.* Weber, *l. c.*, p. 295.

and so that no two consecutive terms of the series

$$1,\ \Delta_1,\ \Delta_2,\ \cdots,\ \Delta_r \tag{39}$$

shall be equal to zero, the last term itself being different from zero. Then the given quadratic function can be expressed as the sum of r *squares. The number of positive squares will be the same as the number of permanences in sign in the series 39 and the number of negative squares will be the same as the number of variations in sign in the same series, provided none of the terms in that series be equal to zero. If, however, one of the terms of the series be equal to zero we shall pay no attention to the two adjacent intervals; to every such pair of intervals there will correspond one positive and one negative square.*

The following theorem is well known :*

If a quadratic function be expressed as a sum of squares of independent linear functions, the number of positive terms is always the same and similarly the number of negative terms is always the same, whatever be the mode of transformation employed.

To complete our discussion we must accordingly show that the linear functions obtained by repeated applications of our two transformations are independent. For this purpose it is sufficient to show that one determinant of order r formed from the coefficients of these linear functions does not reduce to zero † Such a determinant can be formed by taking the coefficients of $x_1, x_2, \cdots, x_r$ in these functions ; for it is easy to see that this determinant is equal to the product of non-vanishing factors of the form $a^{(p)}_{p+1,\,p+1}$ and of the form

$$\begin{vmatrix} a^{(p)}_{p+1,p+2}, & a^{(p)}_{p+1,p+2} \\ a^{(p)}_{p+1,p+2}, & -\,a^{(p)}_{p+1,p+2} \end{vmatrix}.$$

NEW YORK, MARCH 1903.

---

* *Cf.* Weber, *l. c.*, p. 213.

† *Cf.* Weber, *l.* c., p. 108; Bôcher, ANNALS OF MATHEMATICS, vol. 2 (1900–01), p. 84.